

# Comparative Analyses of Single-Nucleotide Polymorphisms in the *TNF* Promoter Region Provide Further Validation for the Vervet Monkey Model of Obesity

Stanton B Gray,<sup>1,2\*</sup> Timothy D Howard,<sup>2</sup> Carl D Langefeld,<sup>3</sup> Gregory A Hawkins,<sup>2</sup> Abdoulaye F Diallo,<sup>2</sup> and Janice D Wagner<sup>1</sup>

Tumor necrosis factor is a cytokine that plays critical roles in inflammation, the innate immune response, and a variety of other physiologic and pathophysiologic processes. In addition, TNF has recently been shown to mediate an intersection of chronic, low-grade inflammation and concurrent metabolic dysregulation associated with obesity and its comorbidities. As part of an ongoing initiative to further characterize vervet monkeys originating from St Kitts as an animal model of obesity and inflammation, we sequenced and genotyped the human ortholog *TNF* gene and approximately 1 kb of the flanking 3' and 5' regions from 265 monkeys in a closed, pedigreed colony. This process revealed a total of 11 single-nucleotide polymorphisms (SNPs) and a single 4-bp insertion–deletion, with minor allele frequencies of 0.08 to 0.39. Many of these polymorphisms were in strong or complete linkage disequilibrium with each other, and all but 1 were contained within a single haplotype block, comprising 5 haplotypes with frequencies of 0.075 to 0.298. Using sequences from humans, chimpanzees, vervets, baboons, and rhesus macaques, phylogenetic shadowing of the *TNF* promoter region revealed that vervet SNPs, like the SNPs in related species, were clustered nonrandomly and nonuniformly around conserved transcription factor binding sites. These data, combined with previously defined heritable phenotypes, permit future association analyses in this nonhuman primate model and have great potential to help dissect the genetic and nongenetic contributions to complex diseases like obesity. More broadly, the sequence data and comparative analyses reported herein facilitates study of the evolution of regulatory sequences of inflammatory and immune-related genes.

**Abbreviations:** NHP, nonhuman primate; SNP, single-nucleotide polymorphism; VRC, Vervet Research Colony.

Obesity is a worldwide epidemic, and a major risk factor for several chronic diseases including type 2 diabetes and cardiovascular disease.<sup>68</sup> Chronic, low-grade, systemic inflammation is associated with obesity and plays an important role in the pathogenesis of type 2 diabetes and cardiovascular disease.<sup>53</sup> Excessive adipose tissue, as seen in overweight and obese states, is a significant source of circulating inflammatory cytokines, which represent a potential pathogenic mechanism linking obesity to these comorbidities. Many cytokine regulators of the putative inflammatory cascade of obesity have been identified. Of these, TNF is among the most well studied. TNF is secreted by adipocytes and other tissues and plays critical roles in lymphocyte biology, immune responses, and promoting inflammation;<sup>32</sup> TNF also increases insulin resistance by downregulating glucose transporter 4 in adipocytes.<sup>30</sup> Serum TNF levels are correlated with obesity and serum levels of C-reactive protein, leptin, and other inflammatory cytokines.<sup>13,53</sup>

Except for some rare monogenic forms, obesity, type 2 diabetes, and cardiovascular disease are classic examples of complex, polygenic diseases. As such, expression of these disease phenotypes requires a combination of genetic susceptibility at multiple gene loci and various environmental risk factors such as lifestyle habits, socioeconomic status, and diet.<sup>28,41,43,44,48,51</sup> In addition, there are significant gender-<sup>45,46</sup> and age-related components to progression and severity of disease expression. Fully delineating these components and their relative contributions has been problematic in human studies.<sup>41</sup> Animal models, of course, provide essential research avenues to control or specifically manipulate environmental factors, thereby yielding more powerful studies through increased homogeneity of conditions, as well as enable the study of disease progression throughout the lifetime of the animal. Old World nonhuman primates (NHPs) are particularly important animal models for the study of obesity and inflammation, given their close phylogenetic relationship to humans, comparable fat distribution,<sup>7</sup> similar reproductive and endocrine physiology to humans, and their potential to exhibit spontaneous obesity, insulin resistance, type 2 diabetes, and cardiovascular disease.<sup>33,34,62–66</sup>

Even though obesity and its comorbidities are polygenic diseases, detailed analysis of each gene implicated is needed to better understand its role in the disease cascade, including gene–gene and gene–environment interactions, and ultimately to expand the range of diagnostic and therapeutic targets. As a

Received: 17 Apr 2009. Revision requested: 29 Jun 2009. Accepted: 04 Oct 2009.

<sup>1</sup>Department of Pathology, Section on Comparative Medicine, <sup>2</sup>Center for Human Genomics, <sup>3</sup>Department of Public Health Sciences, Section on Statistical Genetics and Bioinformatics, Wake Forest University School of Medicine, Winston-Salem, North Carolina.

\*Corresponding author. Email: stgray@wfubmc.edu

prime example, certain risk alleles in humans at single nucleotide polymorphisms (SNPs), primarily within the regulatory promoter–enhancer regions, of the *TNF* gene have various significant associations with obesity and related traits in humans, including cardiovascular disease or type 2 diabetes disease risk, insulin resistance, lipid profiles, and serum or expression levels of adiponectin, C-reactive protein, leptin, and TNF.<sup>19</sup> Despite the multitude of genetic association studies, those demonstrating functional SNPs affecting TNF secretion (for example<sup>50</sup>) or in vitro transcription factor binding<sup>35,36,58,67</sup> are rare. A leading reason for the difficulty in identifying functional SNPs in *TNF*, or anywhere within the major histocompatibility complex, is the strong degree of linkage disequilibrium, or nonrandom association of 2 or more alleles, within this region,<sup>47</sup> yielding extended haplotypes.<sup>10</sup>

As part of an ongoing initiative to further genetically and phenotypically characterize the vervet monkey (*Chlorocebus aethiops sabaues*) originating from St Kitts as a model of obesity, insulin resistance, and type 2 diabetes,<sup>33,34</sup> we sequenced and genotyped the *TNF* gene and its promoter region in 265 monkeys from the pedigreed Vervet Research Colony (VRC) and performed comparative analyses with human and other NHPs commonly used in biomedical research. In addition, we tested our prediction that SNPs in the human ortholog vervet *TNF* promoter region would be distributed nonrandomly around conserved transcription factor binding sites, as shown in humans and other NHPs.<sup>2,38</sup>

## Materials and Methods

**Animals.** All methods for the care and use of nonhuman primates conformed to standards from the US Department of Agriculture and Public Health Service, and all experimental protocols were approved by the Chancellor's Animal Research Committee at the University of California Los Angeles and the Institutional Animal Care and Use Committee for the Department of Veterans Affairs Greater Los Angeles Healthcare System. The study population comprised 265 vervets (*Chlorocebus aethiops sabaues*) from the VRC, which was founded at the University of California Los Angeles in 1975 from 57 (28 male, 29 female) vervets wild-caught from 1975 to 1983 in St Kitts, West Indies.<sup>17</sup> The current population includes approximately 500 descendants, with 24 of the 29 original matrilines in their 5th to 8th generation, and now resides at the Wake Forest University Primate Center. All vervets were verified to be free from *Mycobacterium tuberculosis* (or *bovis*) infection by using intradermal tuberculin testing. All were fed a commercial maintenance primate laboratory chow (Laboratory Diet 5038; Purina, St Louis, MO) ad libitum and supplemented with fresh fruits and vegetables.

**PCR and sequencing.** Genomic DNA was isolated from peripheral blood samples by using standard protocols.<sup>31</sup> For sequencing, PCR primers were generated with the PrimerSelect software (Laser Gene software, DNASTAR, Madison, WI) by using a consensus *TNF* gene sequence created from human, chimpanzee (*Pan troglodytes*), and rhesus monkey (*Macaca mulatta*) data found in the NCBI gene database (www.ncbi.nlm.nih.gov; uc003nui.1, NM001045511, NM001047149, respectively). Primer consensus sequences were created by using the editing feature in Sequencher version 4.6 (Gene Codes, Ann Arbor, MI) and were designed to amplify the entire *TNF* gene plus approximately 1 kb of the 3' and approximately 1 kb of the 5' flanking regions. PCR conditions were as follows: initial denaturation at 95 °C for 10 min, followed by 30 cycles of 96 °C for 1 min, 58 to 60 °C for 1 min and

72 °C for 60 to 90 s, with final elongation for 10 min at 72 °C in a 20-μL reaction mixture. PCR products were purified (QuickStep, Edge Biosystems, Gaithersburg, MD) to remove dNTPs and excess primers which interfere with sequencing reactions. Purified PCR products were sequenced using dye-terminator chemistry (BigDye, ABI, Foster City, CA). Dye-terminator reactions were performed in 5- to 10-μL volumes in 96-well plates. Reagents in the dye-terminator kit were diluted 4-fold by using a 2.5× dilution buffer (50 mM Tris, 10 mM MgCl<sub>2</sub>, pH 9.0). Sequencing reactions were purified by ethanol precipitation prior to loading on an automated sequencer (3730XL DNA Analyzer, ABI) using the POP7 sequencing polymer. The sequences of the primers used are described in Table 1.

**Sequence analyses.** Sequences from the 265 VRC vervets were aligned and genotyped by using Sequencher version 4.6 (Gene Codes). Linkage disequilibrium structure, and frequencies of alleles and haplotypes were determined by using Haploview version 4.0.<sup>4</sup> Sequence alignment scores were generated using ClustalW version 1.7.<sup>52</sup> This vervet sequence (EU626004), and all other sequences except baboon, reported herein are in the NCBI gene database (accession numbers noted previously); the baboon *TNF* promoter sequence has been published elsewhere.<sup>2,26,38</sup> Transcription factor binding sites for NFAT, ETS, κ3, CRE, and SP1<sup>15,18,23,54–57</sup> and κ1 and κ2<sup>23</sup> were crosschecked by using ORegAnno<sup>40</sup> as implemented in the UCSC genome browser (www.genome.ucsc.edu). Conserved regions were identified and illustrated by phylogenetic shadowing<sup>8</sup> by using the web-based program eShadow (http://eshadow.dcode.org/).<sup>42</sup> Only validated (that is demonstrated experimentally) human SNPs (build 128; www.ncbi.nlm.nih.gov) are reported in the promoter sequence comparison; of those, only those with 5% minor allele frequency or greater were used to compare with the eShadow output. Locations of SNPs were compiled from dbSNP for humans (Build 128, www.ncbi.nlm.nih.gov), and reports in the literature for chimpanzee<sup>2</sup> and rhesus macaque;<sup>49</sup> *TNF* SNPs for the baboon were not available at the time of this writing.

**Statistical analyses.** Kolmogorov–Smirnov tests, which evaluate whether the empirical distribution of a sample statistically deviates from a specified theoretical distribution, were performed by using Analyze-It for Microsoft Excel version 2.12 (http://www.analyze-it.com) to test whether the distribution of the *P* values deviated from uniform distribution; under the null hypothesis, the *P* values of our test follow a uniform distribution. Therefore, the distribution of SNPs *P* values along a 50-bp sliding window for each species (human, chimp, vervet, and rhesus) plus the consensus distribution of all 4 were tested for nonuniformity.

## Results

Analysis of approximately 1 kb 5' upstream, approximately 0.9 kb 3' downstream, all 4 exons, and all intronic regions of the human ortholog vervet *TNF* gene in 265 vervet monkeys revealed a total of 12 polymorphisms: 11 SNPs and one 4-bp insertion–deletion, with minor allele frequencies ranging from 0.08 to 0.39 (Table 2). None of the correlating SNPs have been reported to be polymorphic in human populations (dbSNP build 128). The +611 (TGAA)<sub>4/5</sub> insertion–deletion, located in intron 1, is (TGAA)<sub>6</sub> in humans and apparently is not polymorphic; chimpanzees also contain (TGAA)<sub>6</sub>. The rhesus monkey sequence is reported as (TGAA)<sub>5</sub>, but it is unknown whether this sequence is polymorphic in rhesus or any other NHP species. The entire human or-

**Table 1.** Sequences of primers used in this study

5' Forward 3'	5' Reverse 3'
AGA TGA AGA GTG AGA GGG CAT G	AAC GTC CCC TGT ATT CCA TAC C
ACC TCT GGG GAG ATG TGA CC	GAG GGG CTT CAG AAA GCT GA
TTG GAA GCC AAG ACT GAA ACC	CGG GGA TTT GGA AAG TTG G
TTT CCT GCA TCC TGT CTG GA	TCC TTG CTG AGG GAG CGT
GAC AGA AGG TGC AGG GCC	GTG GGA GAG TGG ATG AAG GC
ATG ATC CGG GAC GTG GAG	ACT TGT TTC TTC CCC CAT CTC
CTG CTG CAC TTT GGA GTG AT	TCA CCC TTA AAG GAG GAA CAG
AGA TGG GGG AAG AAA CAA GTG	CAA CAT GGG CTA CAG GCT TG
AGT TTT GGT CTT GGG GGA GG	AGC CTT GGC CCT TGA AGA G
TGG AGA GTG AAC CGA CAT GG	TCC TCC TCA CAG GGC AAT G
CAA GAG CCC CTG CCA GAG	CAA AGG CTC CCT GGT CTC C
TGC ACA GTG AAG TGC TGG C	TGG GGA GCA GAG GCT CAG
CTG AAC AAT AGG CTG TTC CC	CCT GGG GGA TGG GGA ATA T
GAT CAA TCT GCC CGA CTA TCT CG	CCA AAG GCT CCC TGG TCT CC
GAG CCC CTG CCA GAG GGA GAC	TCT CCA GAT TCC AGA TGT CA

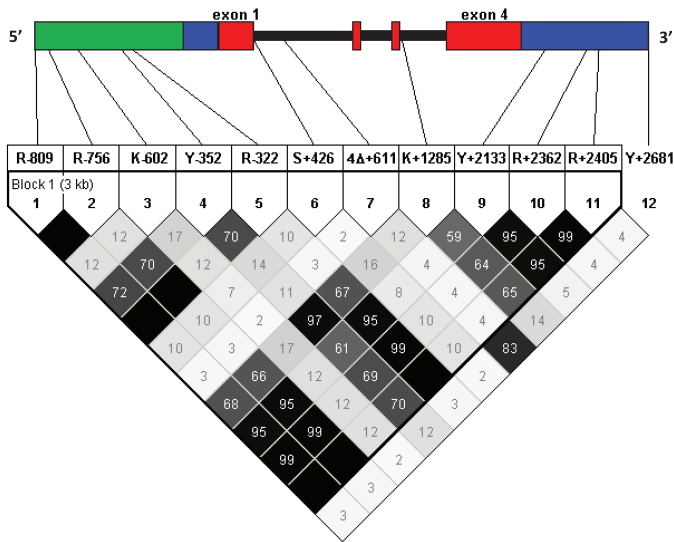
**Table 2.** Polymorphisms in the human ortholog St. Kitts-origin vervet *TNF* gene and flanking sequence.

Polymorphism	Minor allele frequency	Minor allele
A-809G	0.309	G
A-756G	0.311	A
G-602T	0.210	G
C-352T	0.393	C
A-322G	0.312	A
C+426G	0.209	C
+611(tgaa) <sub>4/5</sub>	0.078	(tgaa) <sub>4</sub>
G+1285T	0.387	T
C+2133T	0.309	T
A+2362G	0.316	A
A+2405G	0.318	A
C+2681T	0.077	C

Polymorphism names indicate the positions of the variants relative to the transcriptional start site (+1) of the reference human sequence EU626004. Repeat polymorphisms are denoted by including the repeated sequence in parentheses, with the number of repeat variants subscripted and separated by a slash (for example, (tgaa)<sub>4/5</sub>).

tholog vervet *TNF* gene was contained within a single haplotype block (Figure 1), as determined by the confidence interval algorithm<sup>22</sup> implemented within Haploview 4.0.<sup>4</sup> The C/T+2681 SNP located at the farthest 3' untranslated region was the only SNP not included with this block (Figure 1). Many of the polymorphisms were in strong or complete linkage disequilibrium with each other (Figure 1), as is common within and between genes and intergenic regions of the human major histocompatibility complex.<sup>10,47</sup> Five haplotypes were contained in this block, with frequencies ranging from 0.075 to 0.298 (Table 3).

Sequence identities for the 1176-bp *TNF* promoter region between human and chimpanzee, St Kitts-origin vervet, and rhesus macaque were 99.7%, 94.4%, 94.6%, and 94.5%, respectively (Table 4); identities for the entire approximately 4-kb *TNF* gene including the 5' and 3' flanking regions were greater than 99%, 93.3%, and 93.1%, respectively. The entire *TNF* sequence was not



**Figure 1.** Linkage disequilibrium structure of the St-Kitts-origin vervet *TNF* locus. The single haplotype block (outlined triangle) was determined with the confidence interval algorithm<sup>22</sup> of HaploView. Numbers indicate pairwise  $r^2$  values  $\times 100$ . Darker blocks denote higher correlation, and black squares with no numbers represent complete linkage disequilibrium ( $r^2 = 1$ ). Relative positions of polymorphisms are indicated on diagram of the human ortholog St-Kitts-origin vervet *TNF* gene and flanking sequence. Colored regions denote promoter-enhancer region (green), untranslated regions (blue), exons (red), and introns (black).

**Table 3.** Haplotypes and frequencies of the human ortholog St-Kitts-origin vervet *TNF* gene and flanking sequences

Haplotype	Frequency
G A T C A G (tgaa) <sub>5</sub> T T A A	0.298
A G G T G G (tgaa) <sub>5</sub> G C G G	0.208
A G T T G C (tgaa) <sub>5</sub> G C G G	0.205
A G T T G G (tgaa) <sub>5</sub> G C G G	0.183
A G T C G G (tgaa) <sub>4</sub> T C G G	0.075

**Table 4.** Distance matrices and DNA sequence identity between human, chimpanzee, vervet, and rhesus macaque.

	<i>P. troglodytes</i>		<i>C. aethiops</i>		<i>M. mulatta</i>			
	DM	%ID <sub>tr</sub>	DM	%ID <sub>tr</sub>	DM	%ID <sub>tr</sub>		
<i>H. sapiens</i>	0.00	100	7.13	94.20	7.55	94.0		
<i>P. troglodytes</i>			7.13	94.20	7.55	94.0		
<i>C. aethiops</i>					1.89	98.5		
			DM	%ID <sub>tot</sub>	%ID <sub>pr</sub>	DM	%ID <sub>tot</sub>	%ID <sub>pr</sub>
<i>H. sapiens</i>			7.71	93.3	92.9	7.97	93.1	92.3
<i>C. aethiops</i>						2.12	98.4	97.7

DM, distance matrix; %ID, percentage of DNA sequence identity (tr, DNA transcript; tot, total *TNF* sequence including 3' and 5' untranslated flanking regions; pr, *TNF* promoter–enhancer region). Baboon sequence was not included because it was not available ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) at the time of this writing.

available for baboon at the time of this writing and therefore was not included in these analyses.

Figure 2 illustrates a sequence comparison of the *TNF* promoter region in the 5 primate species. The transcription binding sites NFAT, ETS,  $\kappa$ 3, CRE, and SP1, confirmed to be essential for *TNF* gene regulation<sup>5,18,25,54–57</sup> are noted;  $\kappa$ 1 and  $\kappa$ 2, which bind NF $\kappa$ B proteins but are not essential to *TNF* gene regulation,<sup>23</sup> are included also. Those sites essential for gene regulation show complete conservation among the species examined here. The vervet, rhesus, and baboon all have A instead of G in  $\kappa$ 2 (position –206), but  $\kappa$ 1 is completely conserved among all species examined. SNP locations for human, chimpanzee, rhesus, and the St Kitts-origin vervet relative to the human transcription start site are noted (Figure 2). All fall outside major blocks of conserved regions known to be critical for *TNF* gene regulation (Figure 3 A). In addition, there is nonrandom clustering of SNPs among the 4 species examined at sites shown to contain alleles in humans significantly associated with risk for a variety of inflammatory diseases or disorders of innate immunity (that is, –308, –857, –863; Figure 3 B).<sup>19</sup>

The statistical significance of these SNP clusters was assessed by using the Komolgorov–Smirnov test. The null hypothesis—that the distribution of SNPs in the 1152-nucleotide sequence 5' to the transcription start site does not differ from uniformity—was rejected ( $P < 0.0001$ ) in all species as well as the consensus sequence (Table 5).

## Discussion

We report a total of 5 SNPs within the 5' untranslated and promoter regions of the human ortholog St-Kitts-origin vervet *TNF* gene, with another 6 SNPs and 1 insertion–deletion polymorphism in the intronic and 3' untranslated and flanking regions (Table 2, Figure 1). All polymorphisms except 1 (C/T+2681) form a single haplotype block (Figure 1), which contains 5 haplotypes (Table 3) within the VRC population. To our knowledge, this information represents the most extensive report of SNPs, haplotypes, and linkage disequilibrium of the *TNF* gene and flanking regions for any NHP, as well as the first report of complete *TNF* gene sequence data for the vervet monkey and any other *Chlorocebus* species.

Old World NHPs are extremely valuable animal models for the study of complex, polygenic diseases like obesity and its comorbidities.<sup>11,59,65</sup> Particularly when pedigree and genomic structure is known, NHP models offer great power to dissect genetic from nongenetic components of disease through increased homogene-

ity of conditions (for example, see references 21 and 61) as well as enable study of disease progression throughout the lifetime of the animal, thereby also controlling for age and various epigenetic effects (for example, see reference 14). In addition, comparative sequence analyses utilizing NHPs are an essential component of research directed at identifying functional regions within the human genome<sup>8,12,26,38,60</sup> as well as providing insight into the evolutionary history of transcriptional regulation.<sup>2,24,47</sup>

We have demonstrated that vervet *TNF* promoter SNPs are nonrandomly and nonuniformly distributed outside transcription factor binding sites and other conserved regions. This distribution is consistent with differential rates of accumulation of mutations along the *TNF* promoter region, as shown previously.<sup>1,2,27,38</sup> Those regions not affecting *TNF* transcription will presumably be under neutral selection, whereas there will be high selective pressure to preserve those regions that are essential to transcription factor binding. An unresolved question is whether SNPs around transcription factor binding sites can modulate transcription (or are in linkage disequilibrium with other MHC alleles causing such effects) and whether specific alleles could impart increased survivability for a subject according to the specific immunologic or metabolic challenges of that subject's environment. The concept that SNPs have accumulated independently in NHPs—and confer analogous disease phenotypes and risk—has been proposed by various investigators working with NHPs as models of neuropsychiatric disorders.<sup>3,6,37,39</sup> Here we provide the first suggestion to our knowledge that SNPs in the promoter–enhancer region of *TNF* could have accumulated in NHPs and humans experiencing similar selective pressures over evolutionary time for modulating immune or inflammatory responses. The demonstration that SNPs in humans, chimpanzees, vervets, and rhesus monkeys all cluster together in 2 regions on the *TNF* promoter–enhancer region (Figure 3 B) is particularly interesting. What makes this finding even more curious is that these regions are within 50 nucleotides of 3 important human SNPs (–308, –857, and –863), which have been significantly associated with numerous obesity- and diabetes-related phenotypes.<sup>19</sup> SNPs in these particular regions may modulate transcription in important but subtle ways and have thus accumulated in these regions independently over evolutionary time in these 4 primate species (human, chimpanzee, vervet, rhesus).

Research directed at further elucidating the genetic and molecular mechanisms of obesity and untangling the age- and environment-related factors through more relevant animal models has

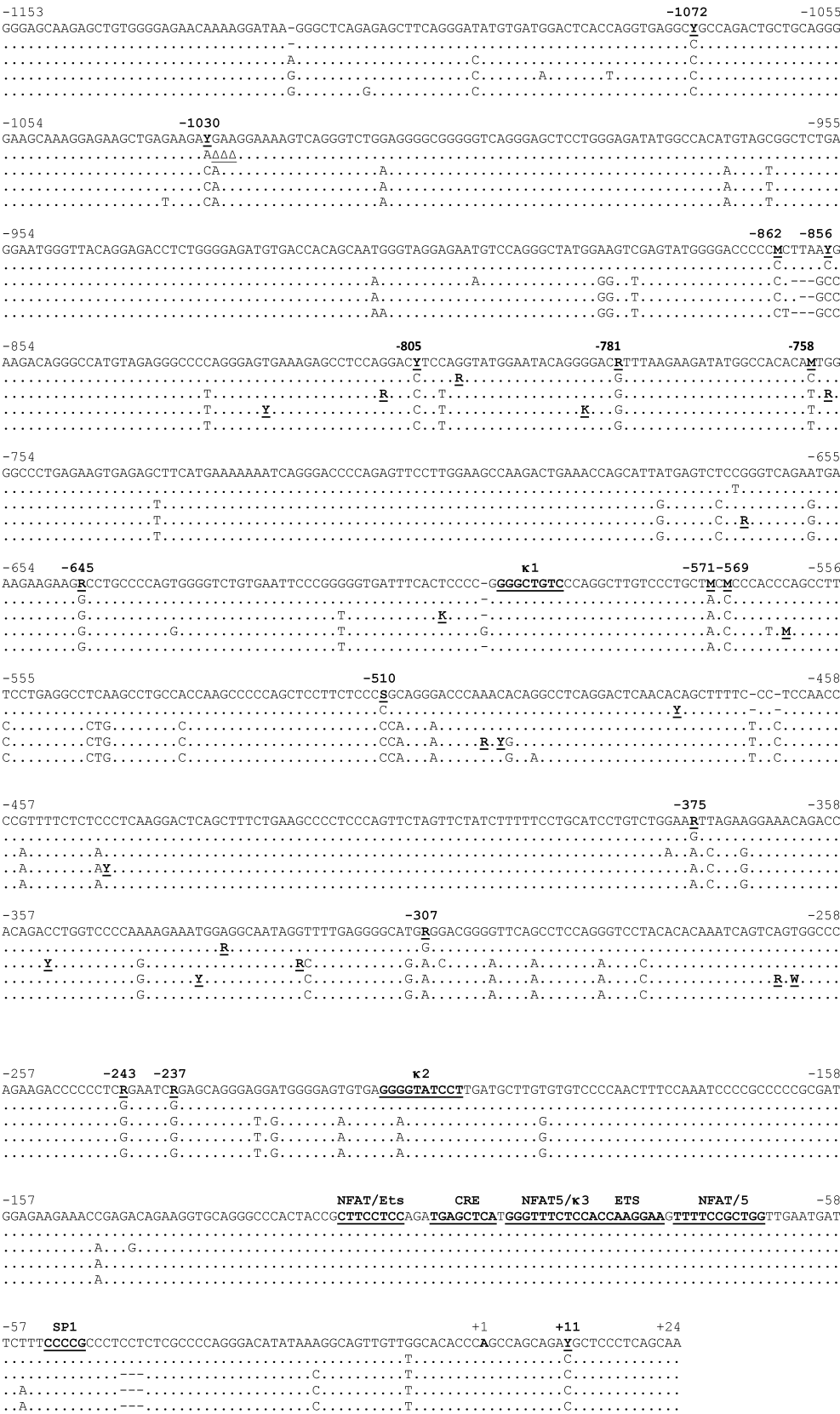
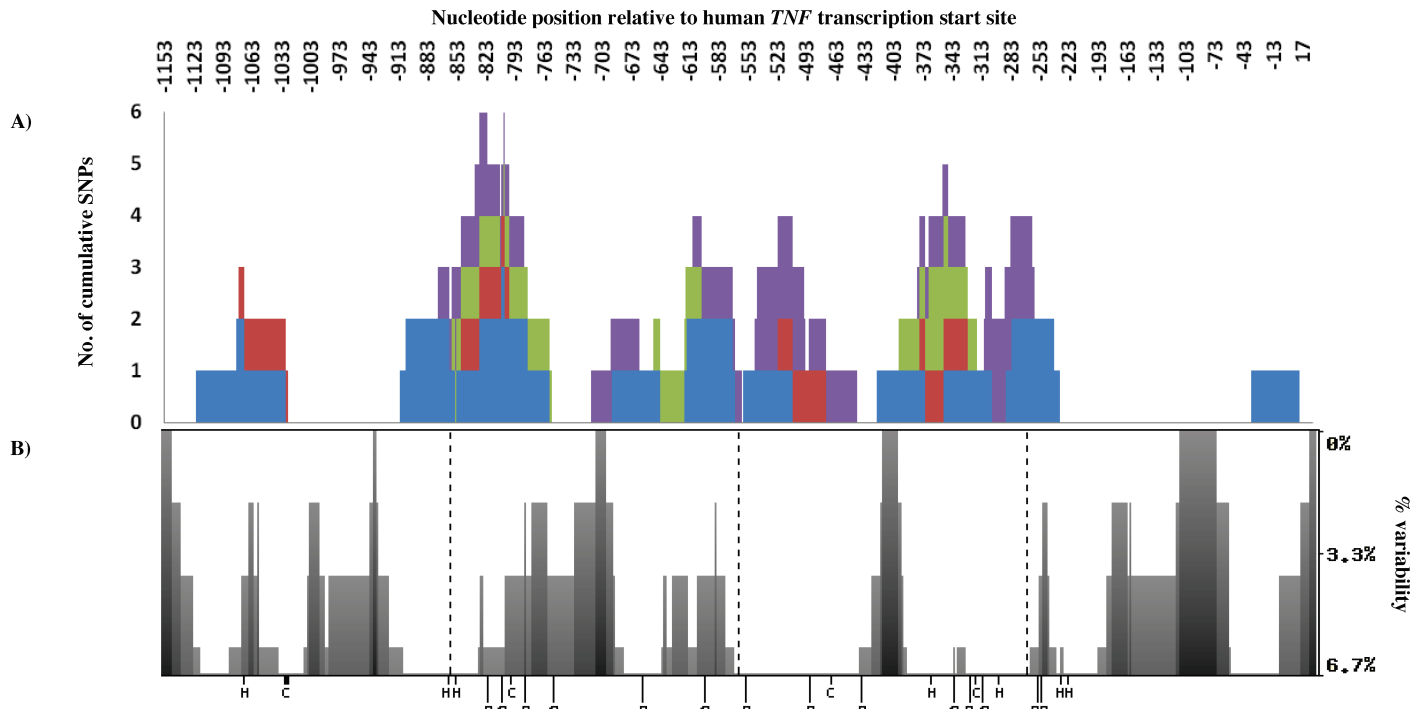


Figure 2. Sequence comparison of *TNF* 5' promoter region between humans (*H. sapiens*), chimpanzee (*P. troglodytes*), St-Kitts-origin vervet (*C. a. sabaeus*), rhesus monkey (*M. mulatta*), and baboon (*P. hamadrayas*); SNPs are underlined and bolded with adjacent nucleotide position for the species in which



**Figure 3.** (A) Phylogenetic shadowing of the *TNF* promoter region in 5 primate species. The multiple-sequence alignment of *TNF* (including the promoter region) for humans, chimpanzees, St-Kitts-origin vervet, rhesus monkey, and baboon was analyzed with eShadow.<sup>42</sup> The top *x* axis is nucleotide position relative to human transcription start site (+1). The relative positions of polymorphisms in 4 of the 5 species studied here are indicated along the bottom *x* axis (H, human; C, chimpanzee; G, St Kitts vervet; R, rhesus monkey); no SNP information for baboon was available. Percentage of nucleotide variation (inverse of conservation) along a 50-nucleotide (nt) sliding window is on the *y* axis, and darker shading of peaks represents regions of higher conservation. (B) Clustering of SNP in NHP. Plot of the number of SNP in each of 4 primate species (human, chimpanzee, vervet monkey, rhesus macaque) which cluster along a sliding 50-nt window of the *TNF* promoter–enhancer region (*y* axis). SNP in each of the 4 species is designated by color (human, blue; chimpanzee, red; vervet, green; rhesus, purple). The *x* axis is nucleotide position relative to the human transcription start site (+1).

**Table 5.** Kolmogorov–Smirnov tests of nonrandom distribution of SNPs

Sequence source	<i>D</i>	<i>p</i>
<i>H. sapien</i>	0.31	<0.0001
<i>P. troglodytes</i>	0.47	<0.0001
<i>C. aethiops sabaues</i>	0.46	<0.0001
<i>M. mulatta</i>	0.41	<0.0001
Consensus	0.19	<0.0001

been assigned as 1 of the highest priorities in the United States.<sup>15</sup> The information presented here represents the second phase of our effort to further genetically characterize the St-Kitts-origin vervet as a model of obesity and type 2 diabetes, both of which involve chronic, low-grade inflammation and concurrent metabolic dysregulation mediated by *TNF*.<sup>9,16,29</sup> This current study adds to prior characterization of obesity-related phenotypic measures and heritability of these traits.<sup>33,34</sup> The vervet monkey has long been 1

of the most important NHP models of biomedical research, with a PubMed citation record over the past 10 y close to that of rhesus macaque and greater than any other NHP.<sup>20</sup> However, gene sequence and genome information for the vervet has begun only recently to emerge.<sup>31</sup> This current report involves 1 of the largest numbers of NHP sequenced for *TNF* or any other individual candidate gene, which has provided sufficient statistical power to uncover both common and relatively rare SNPs and haplotypes as well as to show that the distribution of these polymorphisms is consistent with the regions of conservation or neutral selection in humans and other NHPs. A previous study revealed<sup>2</sup> differential Sp1 transcription factor binding and *TNFs* transcription between various species of Asian apes and all other Old World apes and monkeys, according to nucleotide differences within that transcription factor binding site. However to our knowledge, no such intraspecies comparisons of *TNF* transcription factor binding in NHPs have been attempted. If significant genetic associations are found between obesity-related phenotypes in the vervet and the SNPs or haplotypes presented here, then an important next

they have been reported. Nucleotide positions are relative to the human transcription start site (+1). The single-letter code is used to identify SNPs: M, A or C; R, A or G; W, A or T; S, C or G; Y, C or T; K, G or T. Transcription factor binding sites are underlined, with names above the sequence. NFAT (nuclear factor of activated T cells), Ets,  $\kappa$ 3, CRE (cyclic AMP response element), and SP-1 are critical for *TNF* gene regulation;<sup>5,18,25,54-57</sup>  $\kappa$ 1 and  $\kappa$ 2 are reported to bind NF $\kappa$ B proteins and are not essential for gene regulation.<sup>23,38</sup>

step would be to evaluate those variable nucleotide sequences for differential transcription factor binding. Obesity and obesity-related phenotypes in the VRC have been measured and reported previously,<sup>33</sup> with abdominal obesity correlating well with serum triglyceride levels and insulin resistance, as seen in humans. Genetic association analyses are underway.

The broader value of the analyses we present here addresses the intersection of metabolism and inflammation/innate immunity, which are among the most fundamental requirements for an organism's survival. As such, metabolism and innate immunity show high degrees of evolutionary conservation, are highly integrated, and essentially are governed by a single central homeostatic mechanism,<sup>29</sup> a greater understanding of which is paramount given the current worldwide epidemic of obesity and its comorbidities.<sup>68</sup>

In summary, our report on the sequence of the human ortholog St-Kitts-origin vervet *TNF* gene, as well as its polymorphisms and haplotype and linkage disequilibrium structure, provides both proximate and long-term value into research directed at further elucidating the pathogenic mechanisms of obesity-related inflammation and metabolic dysregulation. Association studies using these SNP and haplotype data are underway and include previously defined obesity-related traits. Identification of any such associations will allow targeted efforts toward understanding the functional relevance of associated polymorphisms through transcription factor binding assays, for example. This approach, in a pedigreed NHP model, offers great potential to dissect genetic from environmental components of polygenic, complex diseases such as obesity and its comorbidities. In addition, the sequence data facilitates future comparative analyses and research into the evolution of regulatory gene sequences of complex disease.

## Acknowledgments

This work was funded in part by the Wake Forest University School of Medicine Venture Fund, the Skorich Diabetes Research Fund, the Monty Blackmon Diabetes Research Fund (JDW), T32 RR07009 from NIH/NICRR (SBG), and an Animal and Biological Materials Resources grant from the National Center for Research Resources (P40 RR 019963 to Dr Lynn Fairbanks at UCLA). Support of this project from the Center for Public Health Genomics (CDL) at Wake Forest University School of Medicine is greatly appreciated. We also thank Drs Lynn Fairbanks (UCLA), Nelson Friemer (UCLA), Matthew Jorgensen (WFUSM), Kylie Kavanagh (WFUSM), and Larry Rudel (WFUSM) for their input and collaboration. This work constitutes partial fulfillment of PhD requirements for Dr Gray.

## References

1. Baena A, Leung JY, Sullivan AD, Landires I, Vasquez-Luna N, Quinones-Berrocal J, Fraser PA, Uko GP, Delgado JC, Clavijo OP, Thim S, Meshnick SR, Nyirenda T, Yunis EJ, Goldfeld AE. 2002. TNF- $\alpha$  promoter single nucleotide polymorphisms are markers of human ancestry. *Genes Immun* 3:482–487.
2. Baena A, Mootnick AR, Falvo JV, Tsytskova AV, Ligeiro F, Diop OM, Brieva C, Gagneux P, O'Brien SJ, Ryder OA, Goldfeld AE. 2007. Primate TNF promoters reveal markers of phylogeny and evolution of innate immunity. *PLoS One* 2:e621.
3. Bailey JN, Breidenthal SE, Jorgensen MJ, McCracken JT, Fairbanks LA. 2007. The association of DRD4 and novelty seeking is found in a nonhuman primate model. *Psychiatr Genet* 17:23–27.
4. Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265.
5. Barthel R, Tsytskova AV, Barczak AK, Tsai EY, Dascher CC, Brenner MB, Goldfeld AE. 2003. Regulation of tumor necrosis factor  $\alpha$  gene expression by mycobacteria involves the assembly of a unique enhanceosome dependent on the coactivator proteins CBP/p300. *Mol Cell Biol* 23:526–533.
6. Bennett AJ, Lesch KP, Heils A, Long JC, Lorenz JG, Shoaf SE, Champoux M, Suomi SJ, Linnoila MV, Higley JD. 2002. Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Mol Psychiatry* 7:118–122.
7. Bergman RN, Kim SP, Hsu IR, Catalano KJ, Chiu JD, Kabir M, Richey JM, Ader M. 2007. Abdominal obesity: role in the pathophysiology of metabolic disease and cardiovascular risk. *Am J Med* 120(2 Suppl 1):S3–S8; discussion S29–S32.
8. Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, Pachter L, Rubin EM. 2003. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* 299:1391–1394.
9. Cawthorn WP, Sethi JK. 2008. TNF $\alpha$  and adipocyte biology. *FEBS Lett* 582:117–131.
10. Ceppellini R, Siniscalco M, Smith CA. 1955. The estimation of gene frequencies in a random-mating population. *Ann Hum Genet* 20:97–115.
11. Comuzzie AG, Cole SA, Martin L, Carey KD, Mahaney MC, Blangero J, VandeBerg JL. 2003. The baboon as a nonhuman primate model for the study of the genetics of obesity. *Obes Res* 11:75–80.
12. Cooper GM, Sidow A. 2003. Genomic regulatory regions: insights from comparative sequence analysis. *Curr Opin Genet Dev* 13:604–610.
13. Coppack SW. 2001. Proinflammatory cytokines and adipose tissue. *Proc Nutr Soc* 60:349–356.
14. de Vries A, Holmes MC, Heijnis A, Seier JV, Heerden J, Louw J, Wolfe-Coote S, Meaney MJ, Levitt NS, Seckl JR. 2007. Prenatal dexamethasone exposure induces changes in nonhuman primate offspring cardiometabolic and hypothalamic-pituitary-adrenal axis function. *J Clin Invest* 117:1058–1067.
15. Eckel RH, Barouch WW, Ershow AG. 2002. Report of the National Heart, Lung, and Blood Institute-National Institute of Diabetes and Digestive and Kidney Diseases Working Group on the pathophysiology of obesity-associated cardiovascular disease. *Circulation* 105:2923–2928.
16. Esteve E, Ricart W, Fernandez-Real JM. 2005. Dyslipidemia and inflammation: an evolutionary conserved mechanism. *Clin Nutr* 24:16–31.
17. Fairbanks LA, Newman TK, Bailey JN, Jorgensen MJ, Breidenthal SE, Ophoff RA, Comuzzie AG, Martin LJ, Rogers J. 2004. Genetic contributions to social impulsivity and aggressiveness in vervet monkeys. *Biol Psychiatry* 55:642–647.
18. Falvo JV, Ugialoro AM, Brinkman BM, Merika M, Parekh BS, Tsai EY, King HC, Morielli AD, Peralta EG, Maniatis T, Thanos D, Goldfeld AE. 2000. Stimulus-specific assembly of enhancer complexes on the tumor necrosis factor  $\alpha$  gene promoter. *Mol Cell Biol* 20:2239–2247.
19. Fernandez-Real JM. 2006. Genetic predispositions to low-grade inflammation and type 2 diabetes. *Diabetes Technol Ther* 8:55–66.
20. Freimer NB, Dewar K, Kaplan JR, Fairbanks LA. [Internet]. The importance of the vervet (African green monkey) as a biomedical model. [Cited 07 March 2008]. Available at <http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/TheVervetMonkeyBiomedicalModel.pdf>.
21. Freimer NB, Service SK, Ophoff RA, Jasinska AJ, McKee K, Ville-neuve A, Belisle A, Bailey JN, Breidenthal SE, Jorgensen MJ, Mann JJ, Cantor RM, Dewar K, Fairbanks LA. 2007. A quantitative trait locus for variation in dopamine metabolism mapped in a primate model using reference sequences from related species. *Proc Natl Acad Sci USA* 104:15811–15816.
22. Gabriel SB, Schaffner SE, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ,

- Altshuler D. 2002. The structure of haplotype blocks in the human genome. *Science* **296**:2225–2229.
23. Goldfeld AE, Doyle C, Maniatis T. 1990. Human tumor necrosis factor  $\alpha$  gene regulation by virus and lipopolysaccharide. *Proc Natl Acad Sci USA* **87**:9769–9773.
24. Goldfeld AE, Leung JY, Sawyer SA, Hartl DL. 2000. Post-genomics and the neutral theory: variation and conservation in the tumor necrosis factor  $\alpha$  promoter. *Gene* **261**:19–25.
25. Goldfeld AE, McCaffrey PG, Strominger JL, Rao A. 1993. Identification of a novel cyclosporin-sensitive element in the human tumor necrosis factor  $\alpha$  gene promoter. *J Exp Med* **178**:1365–1379.
26. Haudek SB, Natmessnig BE, Redl H, Schlag G, Giroir BP. 1998. Genetic sequences and transcriptional regulation of the TNF $\alpha$  promoter: comparison of human and baboon. *Immunogenetics* **48**:202–207.
27. Higasa K, Hayashi K. 2006. Periodicity of SNP distribution around transcription start sites. *BMC Genomics* **7**:66.
28. Hill JO, Wyatt HR, Reed GW, Peters JC. 2003. Obesity and the environment: where do we go from here? *Science* **299**:853–855.
29. Hotamisligil GS. 2006. Inflammation and metabolic disorders. *Nature* **444**:860–867.
30. Hotamisligil GS, Shargill NS, Spiegelman BM. 1993. Adipose expression of tumor necrosis factor- $\alpha$ : direct role in obesity-linked insulin resistance. *Science* **259**:87–91.
31. Jasinska AJ, Service S, Levinson M, Slaten E, Lee O, Sobel E, Fairbanks LA, Bailey JN, Jorgensen MJ, Breidenthal SE, Dewar K, Hudson TJ, Palmour R, Freimer NB, Ophoff RA. 2007. A genetic linkage map of the vervet monkey (*Chlorocebus aethiops sabaeus*). *Mamm Genome* **18**:347–360.
32. Juge-Aubry CE, Henrichot E, Meier CA. 2005. Adipose tissue: a regulator of inflammation. *Best Pract Res Clin Endocrinol Metab* **19**:547–566.
33. Kavanagh K, Fairbanks LA, Bailey JN, Jorgensen MJ, Wilson M, Zhang L, Rudel LL, Wagner JD. 2007. Characterization and heritability of obesity and associated risk factors in vervet monkeys. *Obesity (Silver Spring)* **15**:1666–1674.
34. Kavanagh K, Jones KL, Sawyer J, Kelley K, Carr JJ, Wagner JD, Rudel LL. 2007. Trans fat diet induces abdominal obesity and changes in insulin sensitivity in monkeys. *Obesity (Silver Spring)* **15**:1675–1684.
35. Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, Marsh K, Kwiatkowski D. 1999. A polymorphism that affects OCT1 binding to the TNF promoter region is associated with severe malaria. *Nat Genet* **22**:145–150.
36. Kroeger KM, Carville KS, Abraham LJ. 1997. The -308 tumor necrosis factor $\alpha$  promoter polymorphism effects transcription. *Mol Immunol* **34**:391–399.
37. Lesch KP, Meyer J, Glatz K, Flugge G, Hinney A, Hebebrand J, Klauk SM, Poustka A, Poustka F, Bengel D, Mossner R, Riederer P, Heils A. 1997. The 5HT transporter gene-linked polymorphic region (5HTTLPR) in evolutionary perspective: alternative biallelic variation in rhesus monkeys. *Rapid communication. J Neural Transm* **104**:1259–1266.
38. Leung JY, McKenzie FE, Uglialoro AM, Flores-Villanueva PO, Sorkin BC, Yunis EJ, Hartl DL, Goldfeld AE. 2000. Identification of phylogenetic footprints in primate tumor necrosis factor  $\alpha$  promoters. *Proc Natl Acad Sci USA* **97**:6614–6618.
39. Miller GM, Bendor J, Tiefenbacher S, Yang H, Novak MA, Madras BK. 2004. A  $\mu$ -opioid receptor single nucleotide polymorphism in rhesus monkey: association with stress response and aggression. *Mol Psychiatry* **9**:99–108.
40. Montgomery SB, Griffith OL, Sleumer MC, Bergman CM, Bilenky M, Pleasance ED, Prychyna Y, Zhang X, Jones SJ. 2006. ORegAnno: an open access database and curation system for literature-derived promoters, transcription factor binding sites and regulatory variation. *Bioinformatics* **22**:637–640.
41. Mutch DM, Clement K. 2006. Unraveling the genetics of human obesity. *PLoS Genet* **2**:e188.
42. Ovcharenko I, Boffelli D, Loots GG. 2004. eShadow: a tool for comparing closely related sequences. *Genome Res* **14**:1191–1198.
43. Pasquali R, Vicennati V, Cacciari M, Pagotto U. 2006. The hypothalamic-pituitary-adrenal axis activity in obesity and the metabolic syndrome. *Ann N Y Acad Sci* **1083**:111–128.
44. Roseboom TJ, van der Meulen JH, Ravelli AC, Osmond C, Barker DJ, Bleker OP. 2001. Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview. *Mol Cell Endocrinol* **185**:93–98.
45. Salpeter SR, Walsh JM, Ormiston TM, Greyber E, Buckley NS, Salpeter EE. 2006. Meta-analysis: effect of hormone-replacement therapy on components of the metabolic syndrome in postmenopausal women. *Diabetes Obes Metab* **8**:538–554.
46. Schneider JG, Tompkins C, Blumenthal RS, Mora S. 2006. The metabolic syndrome in women. *Cardiol Rev* **14**:286–291.
47. Shiina T, Ota M, Shimizu S, Katsuyama Y, Hashimoto N, Takasu M, Anzai T, Kulski JK, Kikkawa E, Naruse T, Kimura N, Yanagiya K, Watanabe A, Hosomichi K, Kohara S, Iwamoto C, Umehara Y, Meyer A, Wanner V, Sano K, Macquin C, Ikeo K, Tokunaga K, Gojobori T, Inoko H, Bahram S. 2006. Rapid evolution of major histocompatibility complex class I genes in primates generates new disease alleles in humans via hitchhiking diversity. *Genetics* **173**:1555–1570.
48. Shively CA, Clarkson TB. 1988. Regional obesity and coronary artery atherosclerosis in females: a nonhuman primate model. *Acta Med Scand Suppl* **723**:71–78.
49. Singh KK, Schmidtke J. 2005. Single nucleotide polymorphisms within the promoter region of the rhesus monkey tumor necrosis factor  $\alpha$  gene. *Immunogenetics* **57**:289–292.
50. Skoog T, Eriksson P, Hoffstedt J, Ryden M, Hamsten A, Arner P. 2001. Tumour necrosis factor  $\alpha$  (TNF $\alpha$ ) polymorphisms-857C/A and -863C/A are associated with TNF $\alpha$  secretion from human adipose tissue. *Diabetologia* **44**:654–655.
51. Taylor PD, Poston L. 2007. Developmental programming of obesity in mammals. *Exp Physiol* **92**:287–298.
52. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res* **22**:4673–4680.
53. Trayhurn P, Wood IS. 2004. Adipokines: inflammation and the pleiotropic role of white adipose tissue. *Br J Nutr* **92**:347–355.
54. Tsai EY, Falvo JV, Tsytsykova AV, Barczak AK, Reimold AM, Glimcher LH, Fenton MJ, Gordon DC, Dunn IF, Goldfeld AE. 2000. A lipopolysaccharide-specific enhancer complex involving Ets, Elk1, Sp1, and CREB binding protein and p300 is recruited to the tumor necrosis factor  $\alpha$  promoter in vivo. *Mol Cell Biol* **20**:6084–6094.
55. Tsai EY, Jain J, Pesavento PA, Rao A, Goldfeld AE. 1996. Tumor necrosis factor  $\alpha$  gene regulation in activated T cells involves ATF-2/Jun and NFATp. *Mol Cell Biol* **16**:459–467.
56. Tsai EY, Yie J, Thanos D, Goldfeld AE. 1996. Cell-type-specific regulation of the human tumor necrosis factor  $\alpha$  gene in B cells and T cells by NFATp and ATF-2/JUN. *Mol Cell Biol* **16**:5232–5244.
57. Tsytsykova AV, Goldfeld AE. 2002. Inducer-specific enhanceosome formation controls tumor necrosis factor  $\alpha$  gene expression in T lymphocytes. *Mol Cell Biol* **22**:2620–2631.
58. Udalova IA, Richardson A, Denys A, Smith C, Ackerman H, Foxwell B, Kwiatkowski D. 2000. Functional consequences of a polymorphism affecting NF $\kappa$ B p50-p50 binding to the TNF promoter region. *Mol Cell Biol* **20**:9113–9119.
59. VandeBerg JL, Williams-Blangero S. 2003. International perspectives: the future of nonhuman primate resources: proceedings of the workshop held 17–19 Apr 2002. Washington (DC): National Academies Press.
60. Villinger F, Brar SS, Mayne A, Chikkala N, Ansari AA. 1995. Comparative sequence analysis of cytokine genes from human and nonhuman primates. *J Immunol* **155**:3946–3954.
61. Vinson A, Mahaney MC, Cox LA, Rogers J, VandeBerg JL, Rainwater DL. 2008. A pleiotropic QTL on 2p influences serum LpPLA2

- activity and LDL cholesterol concentration in a baboon model for the genetics of atherosclerosis risk factors. *Atherosclerosis* **196**:667–673.
62. **Wagner JD, Bagdade JD, Litwak KN, Zhang L, Bell-Farrow AD, Wang ZQ, Cefalu WT.** 1996. Increased glycation of plasma lipoproteins in diabetic cynomolgus monkeys. *Lab Anim Sci* **46**:31–35.
63. **Wagner JD, Carlson CS, O'Brien TD, Anthony MS, Bullock BC, Cefalu WT.** 1996. Diabetes mellitus and islet amyloidosis in cynomolgus monkeys. *Lab Anim Sci* **46**:36–41.
64. **Wagner JD, Cline JM, Shadoan MK, Bullock BC, Rankin SE, Cefalu WT.** 2001. Naturally occurring and experimental diabetes in cynomolgus monkeys: a comparison of carbohydrate and lipid metabolism and islet pathology. *Toxicol Pathol* **29**:142–148.
65. **Wagner JE, Kavanagh K, Ward GM, Auerbach BJ, Harwood HJ Jr, Kaplan JR.** 2006. Old world nonhuman primate models of type 2 diabetes mellitus. *ILAR J* **47**:259–271.
66. **West DB, York B.** 1998. Dietary fat, genetic predisposition, and obesity: lessons from animal models. *Am J Clin Nutr* **67**:505S–512S.
67. **Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW.** 1997. Effects of a polymorphism in the human tumor necrosis factor  $\alpha$  promoter on transcriptional activation. *Proc Natl Acad Sci USA* **94**:3195–3199.
68. **Wyatt SB, Winters KP, Dubbert PM.** 2006. Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Am J Med Sci* **331**:166–174.